



**UTS**

**UNIVERSITY  
OF TECHNOLOGY  
SYDNEY**

Faculty of Engineering & Information Technology

**Intelligent Personalized Approaches for Semantic Search and Query  
Expansion**

**by**

Omar Ghaleb Alshaweeh

A Thesis Submitted for the Degree of Doctor of Philosophy

February 22, 2019

## **CERTIFICATE**

*I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.*

*I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.*

*This research is supported by the Australian Government Research Training Program.*

*Signature of Student:*

Production Note:  
Signature removed prior to publication.

*Date:*

*22/2/2019*

## **ACKNOWLEDGEMENTS**

First and foremost, my sincere thanks go to Allah, who endowed me to complete this doctorate. In particular, I am grateful to Al-Hussein bin Talal University, Ma'an, Jordan for their financial support of this project.

Most of all, I wish to thank my supervisory Dr. Farookh Khadeer Hussain and Dr. Hai Yan for giving me the opportunity to perform this work and having guided and helped me throughout the research. Their assistance and advice have made this a rewarding experience.

I am extremely grateful to my mother, father, brothers and sisters for all of the sacrifices that you've made on my behalf. Your prayers for me have sustained me thus far. I will never be able to pay back the love and affection showered upon me by my family. I especially wish to thank my wife, Layla, who has been extremely supportive of me throughout this entire process and has made countless sacrifices to help me get to this point.

Finally, I would like to give my special thanks to my great friends. Their motivation and continuous support have helped make this research happen and a more than enjoyable experience. I am really very grateful for all you have done for me.

## ABSTRACT

In today's highly advanced technological world, the Internet has taken over all aspects of human life. Many services are advertised and provided to the users through online channels. The user looks for services and obtains them through different search engines. To obtain the best results that meet the needs and requirements of the users, researchers have extensively studied methods such as different personalization methods by which to improve the performance and efficiency of the retrieval process. A key part of the personalization process is the generation of user models. The most commonly used user models are still rather simplistic, representing the user as a vector of ratings or using a set of keywords. Recently, semantic techniques have had a significant importance in the field of personalized querying and personalized web search engines. This thesis focuses on both processes of personalized web search engines, first the reformulation of queries and second ranking query results.

The importance of personalized web search lies in its ability to identify users' interests based on their personal profiles. This work contributes to personalized web search services in three aspects. These contributions can be summarized as follows:

First, it creates user profiles based on a user's browsing behaviour, as well as the semantic knowledge of a domain ontology, aiming to improve the quality of the search results. However, it is not easy to acquire personalized web search results, hence one of the problems that is encountered in this approach is how to get a precise representation of the user interests, as well as how to use it to find search results. The second contribution builds on the first contribution. A personalized web search approach is introduced by integrating user context history into the information retrieval process. This integration process aims to provide search results that meet the user's needs. It also aims to create contextual profiles for the user based on several basic factors: user temporal behaviour during browsing, semantic knowledge of a specific domain ontology, as well as an algorithm based on re-ranking the search results.

The previous solutions were related to the re-ranking of the returned search results to match the user's requirements. The third contribution includes a comparison of three-term weight methods in personalized query expansion. This model has been built to

incorporate both latent semantics and weighting terms. Experiments conducted in the real world to evaluate the proposed personalized web search approach; show promising results in the quality of reformulation and re-ranking processes compared to Google engine techniques.

## TABLE OF CONTENTS

Certificate	ii
Acknowledgements	iii
Abstract	iv
Table Of Contents	vi
List Of Figures	xii
List Of Table	xiv
LIST OF PUBLICATIONS	xv
Chapter 1	17
Introduction	17
1.1 Introduction	17
1.2 Overview	17
1.3 Personalized Search Using a Semantic Search and Ontology Solution	18
1.4 Motivation	19
1.5 Research Objectives	21
1.6 Research Contributions	21
1.7 Scope of the Work	22
1.8 Plan of the Thesis	23
1.9 Conclusion	24
CHAPTER 2	25
Literature Review	25
2.1 Introduction	25
2.2. Information Retrieval	26
2.3 Web Personalization	27

2.3.1 Overview	27
2.3.2. Personalization Definition and Benefits	27
2.3.3 Web Personalization Process	29
2.4 User Profile Acquisition	29
2.4.1 Overview	29
2.4.2 User Information Collection	30
2.4.3 Explicit User Information Collection	30
2.4.4 Implicit User Information Collection	31
2.5 Profile Representation Approaches	32
2.5.1. Overview	32
2.5.2 Weighted Keyword Vectors	32
2.5.3 Semantic Networks	33
2.5.4 Weighted Concepts	33
2.6 Re-ranking Search Results	33
2.6.1 Overview	33
2.6.2 Ranking Using Social Relevance	34
2.6.3 Personalized Ranking for search results	34
2.6.4 Re-ranking Based on An Ontological User Profile	35
2.6.4.1. Ontology	35
2.6.4.2 Ontology Definition and Types	36
2.6.4.3 Ontology Learning Process	37
2.6.4.4 Personalized ontologies	38
2.6.4.5 Re-ranking based on Ontological profile	39
2.7 Contextual Web Search	41
2.8 Query Expansion	45

2.8.1 Overview	45
2.8.2. Query Expansion Using Relevance Feedback	46
2.8.3 Query Expansion Using Grammatically Related Terms	47
2.8.4 Query Expansion Using Ontologies and Thesauruses	48
2.8.5 Query Expansion Using Contextual Information	48
2.8.6 Query Expansion Using Query-logs	49
2.8.9 Evaluation of Query Expansion Techniques	50
2.9 Summary and Conclusion	52
<b>CHAPTER 3</b>	<b>54</b>
<b>PROBLEM DEFINITION</b>	<b>54</b>
3.1 Introduction	54
3.2 Key Concepts	55
3.2.1 Web Search Engines	55
3.2.2 Personalized Web Search	56
3.2.3 Search Engine Evaluation	56
3.2.4 Information Retrieval	56
3.2.5 Ontologies	56
3.2.6 Query Reformulation	57
3.2.7 Query expansion	57
3.2.8 Personalized Query Expansion	57
3.2.9 WordNet	58
3.2.10 Semantics	58
3.2.11 Service Metadata	58
3.2.12 Fuzzy Logic	58
3.2.13. User context	58



3.2.14. Context search	59
3.2.15. Category hierarchy	59
3.3 Problem Overview and Problem Definition	59
3.4 Research Questions	61
3.5 Aim of the research	62
3.6 Research Objectives	62
3.7 Research Approach to Problem Solving:	62
3.7.1 Existing Research Methods	63
3.7.2 Choice of Science and Engineering-based Research Method	63
3.8 Conclusion	66
<b>CHAPTER 4</b>	<b>67</b>
<b>SOLUTION OVERVIEW</b>	<b>67</b>
4.1 Introduction	67
4.2 Overview of the Solution for Personalized Web Search	68
4.3 Overview of the Solution for Personalized Web Search Based on Ontological User Profile	68
4.4 Overview of the Solution For Context-Aware Personalized Web Search Using Navigation History	70
4.5 Overview of the Solution for Personalized Query Expansion	72
4.6 Overview of the Solution for Validation of the proposed Methodology for Semantic-based information Retrieval	74
4.7 Conclusion	75
<b>Chapter 5</b>	<b>77</b>
<b>Personalized Web Search Based on Ontological User Profile</b>	<b>77</b>
5.1. Introduction	77
5.2 Existing Personalized Web Search Approaches	78

5.3 Personalized Web Search Based on Ontological User Profile Methodology	79
5.4 Building an Ontology-Based User Profile	81
5.5 Search Personalization – Re-ranking Algorithm	85
5.6 Experiment Summary	86
5.6.1 Description of the Data Set Used For Experiments	86
5.6.2 Average Precision of top-n Results	87
5.6.3 The Impact of User Behavior On Precision	88
5.7 Experiment Evaluation and Benchmarking	88
5.7.1 The Open Directory Project (ODP)	89
5.7.2 Samen’s Approach	89
5.7.3 Comparison Results and Evaluation of the Accuracy of Re-ranking Approaches	90
5.8 Conclusion	95
<b>Chapter 6</b>	<b>96</b>
<b>Context-Aware Personalized Web Search</b>	<b>96</b>
6.1. Introduction	96
6.2. Contextual personalized web search approaches	97
6.3. Contextual personalized web search methodology	98
6.4 Building an Ontology-Based User Contextual profile	100
6.5 Personalized Search results using a Re-ranking algorithm	105
6.6 Experiment Description	106
6.7 Conclusions	109
<b>Chapter 7</b>	<b>111</b>
<b>Personalized Query Expansion</b>	<b>111</b>
7.1. Introduction	111

7.2. Challenges of query expansion	112
7.3. Personalized Query Expansion techniques	113
7.4. Personalized Query Expansion Methodology	115
7.5. Experiments	124
7.6 Conclusions	132
<b>CHAPTER 8</b>	<b>133</b>
<b>CONCLUSION</b>	<b>133</b>
8.1 Introduction	133
8.2 Problems Addressed in this thesis	133
8.3 Contributions of this thesis	134
8.3.1 Contributions 1: Building an ontological user profile approach for a personalized search	135
8.3.2 Contributions 2: A contextual user profile based on navigation history	135
8.3.3 Contributions 3: A new technique for personal query expansion	136
8.4 Conclusion and Future Work	136
<b>FUTURE WORK</b>	<b>138</b>

## LIST OF FIGURES

Figure 2.1: The organization of the literature review in this thesis.....	26
Figure 2.2: An overview of the literature review of web personalization.....	27
Figure 2.3: An overview of the literature review of the user profile acquisition.....	30
Figure 2.4: An overview of the literature review of the profile representation approaches. ....	32
Figure 2.5: An overview of the literature review on re-ranking.....	34
Figure 2.6: Ontology Learning Process Cycle.....	37
Figure 2.7: An overview of the literature review of query expansion.....	46
Figure 3.1 A science and engineering-based research method.....	64
Figure 4.1 The phases where the user's profile affects the personalization process during the retrieval process.....	67
Figure 4.2 Stepwise process of building an ontological user profile.....	70
Figure 5.1: Steps of the Personalized Web Search Based on Ontological User Profile Methodology.....	80
Figure 5.2: Average precision for top-n results.....	87
Figure 5.3: The effect of user behavior on precision.....	88
Figure 5.4: Comparison of average precision for top-n results for the three approaches .....	94
Figure 6.1: Flowchart of procedures for the proposed approach.....	100
Figure 6.2: Concept values based on user context history.....	101
Figure 6.3: Matrix of user-concept interest score.....	104
Figure 6.4: Average precision for top-n results.....	108
Figure 6.5: The effect of user context on precision.....	109
Figure 7.1: Basics of query expansion.....	112
Figure 7.2: Example of a user query with pre-processing steps.....	117
Figure 7.3: The mother matrix for documents and terms.....	118
Figure 7.4: A term-text matrix based on SVD.....	121
Figure 7.5: The similarity between each pair of document vectors.....	123
Figure 7.6: A simplified flowchart of our proposed model of the query expansion process.....	123

Figure 7.7: F1-score for short texts.....	131
Figure 7.8: F1-score for long.....	131

## LIST OF TABLES

<b>Table 2.1.</b> Comparative analysis of personalized web search based on ontological user profile .....	40
<b>Table 2.2.</b> Comparative analysis of personalized web search based on contextual user profile.....	44
<b>Table 5.1.</b> Concept weights based on user interests .....	84
<b>Table 5.2:</b> Results of the proposed approach.....	91
<b>Table 5.3:</b> Results of Samen’s approach.....	92
<b>Table 5.4:</b> Results of Google’s approach.....	93
<b>Table 5.5:</b> Average precision for top-n results for the three approaches.....	94
<b>Table 7.1:</b> Our dataset description.....	116
<b>Table 7.2:</b> Results of experiment without query expansion using TF.....	125
<b>Table 7.3:</b> Results of experiment without query expansion using TF-IDF.....	126
<b>Table 7.4:</b> Results of experiment without query expansion using TF-IDF-CF....	127
<b>Table 7.5:</b> Results of experiment with query expansion using TF.....	128
<b>Table 7.6:</b> Results of experiment with query expansion using TF-IDF.....	129
<b>Table 7.7:</b> Results of experiment with query expansion using TF-IDF-CF.....	130

## LIST OF ALGORITHM

<b>Algorithm 5.1:</b> Re-ranking of the retrieved search results using an ontological user profile.....	85
<b>Algorithm 6.1:</b> Re-ranking of the retrieved search results using contextual user profile.....	105
<b>Algorithm 7.1:</b> Feature extraction .....	118

---

---

## LIST OF PUBLICATIONS

### Journal Papers

Elshaweesh , O., Lu,H., Alwedyan ,M ., Chebil ,W, 'Context-Aware Personalized Web Search Using Navigation History'. International Journal on Semantic Web and Information Systems (IJSWIS), Submitted.

### Conference Proceeding

ElShaweesh, O., Hussain, F.K., Lu, H., Al-Hassan, M. & Kharazmi, S. 2017, 'Personalized Web Search Based on Ontological User Profile in Transportation Domain', *International Conference on Neural Information Processing*, Springer, pp. 239-48.